CD4 + T-cell epitope prediction using antigen processing constraints

Ramgopal R. Mettu, Tysheena Charles, Samuel J. Landry

PII: S0022-1759(16)30028-X
DOI: doi: 10.1016/j.jim.2016.02.013
Reference: JIM 12151

To appear in: *Journal of Immunological Methods*

Received date: 15 May 2015
Revised date: 10 December 2015
Accepted date: 11 February 2016

Please cite this article as: Mettu, Ramgopal R., Charles, Tysheena, Landry, Samuel J., CD4 + T-cell epitope prediction using antigen processing constraints, *Journal of Immunological Methods* (2016), doi: 10.1016/j.jim.2016.02.013

CD4+ T-cell Epitope Prediction Using Antigen Processing Constraints

Ramgopal R. Mettu[1,*], Tysheena Charles[2], Samuel J. Landry[2]
1 Department of Computer Science and Vector-Borne Infectious Diseases Research Center, Tulane University, New Orleans, LA, USA
2 Department of Biochemistry, Tulane Medical School, New Orleans, LA, USA
∗ **E-mail: rmettu@tulane.edu**

Abstract

T-cell CD4+ epitopes are important targets of immunity against infectious diseases and cancer. State-of-the-art methods for MHC class II epitope prediction rely on supervised learning methods in which an implicit or explicit model of sequence specificity is constructed using a training set of peptides with experimentally tested MHC class II binding affinity.

In this paper we present a novel method for CD4+ T-cell eptitope prediction based on modeling antigen-processing constraints. Previous work indicates that dominant CD4+ T-cell epitopes tend to occur adjacent to sites of initial proteolytic cleavage. Given an antigen with known three-dimensional structure, our algorithm first aggregates four types of conformational stability data in order to construct a profile of stability that allows us to identify regions of the protein that are most accessible to proteolysis. Using this profile, we then construct a profile of epitope likelihood based on the pattern of transitions from unstable to stable regions. We validate our method using 35 datasets of experimentally measured CD4+ T cell responses of mice bearing I-Ab or HLA-DR4 alleles as well as of human subjects.

Overall, our results show that antigen processing constraints provide a significant source of predictive power. For epitope prediction in single-allele systems, our approach can be combined with sequence-based methods, or used in instances where little or no training data is available. In multiple-allele systems, sequence-based methods can only be used if the allele distribution of a population is known. In contrast, our approach does not make use of MHC binding prediction, and is thus agnostic to MHC class II genotypes.

Introduction

Epitope-specific CD4+ T cells have been observed to correlate with protection against infections and cancer [58, 62]; in some cases immunization with single epitope peptides were protective [28, 40, 45]. However, immunization with CD4+ epitope peptides has also been shown to cause immunopathology and death [4, 54]. These studies emphasize the critical need for further analysis of CD4+ T-cell responses, which would be greatly facilitated by accurate epitope prediction.

Endogenous proteins, such as self proteins and viral proteins, are processed in the cytosol and transported into the ER and loaded onto class I MHC (MHCI) molecules. Exogenous proteins are taken up by endo/phagocytosis and processed into peptides and loaded onto MHC class II (MHCII) molecules. MHCI-peptide complexes bind to specific T-cell receptors on CD8+ T cells, which are cytotoxic, while MHCII-peptide complexes bind T-cell receptors on CD4+ T cells, which are more varied in nature. CD4+ T cells provide numerous protective functions as part of the adaptive immune response, including cytokine-mediated and contact-mediated signals to B cells, CD8+ T cells, and innate-immune cells, as well as direct modes of attack on pathogenic agents. While MHCI and MHCII molecules have scores of alleles, their three-dimensional structures are highly conserved; allele variation occurs primarily in the peptide binding groove and influences antigen peptide specificity. The closed binding grooves of MHCI molecules exhibit a preference for 8- to 11-mers, while the open binding grooves of MHCII molecules are less specific, with bound peptides being between 10 and 30 amino acids long.